

SEQUENCE LISTING

(1) GENERAL INFORMATION:

10400 (i) APPLICANT: LI, HAODONG
ADAMS, MARK D

(ii) TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
(B) STREET: 9410 KEY WEST AVENUE
(C) CITY: ROCKVILLE
(D) STATE: MD
(E) COUNTRY: US
(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/348,815
(B) FILING DATE: 08-JUL-1999
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: JONATHAN L. KLEIN
(B) REGISTRATION NUMBER: 41,119
(C) REFERENCE/DOCKET NUMBER: PF126P1D1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 301-309-8504
(B) TELEFAX: 301-309-8439

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1146 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1146

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG AGC TCC CGC ATC GCC AGG GCG CTC GCC TTA GTC GTC ACC CTT CTC 48
Met Ser Ser Arg Ile Ala Arg Ala Leu Ala Leu Val Val Thr Leu Leu

1					5					10					15	
CAC	TTG	ACC	AGG	CTG	GCG	CTC	TCC	ACC	TGC	CCC	GCT	GCC	TGC	CAC	TGC	96
His	Leu	Thr	Arg	Leu	Ala	Leu	Ser	Thr	Cys	Pro	Ala	Ala	Cys	His	Cys	
			20					25					30			
CCC	CTG	GAG	GCG	CCC	AAG	TGC	GCG	CCG	GGA	GTC	GGG	CTG	GTC	CGG	GAC	144
Pro	Leu	Glu	Ala	Pro	Lys	Cys	Ala	Pro	Gly	Val	Gly	Leu	Val	Arg	Asp	
		35					40					45				
GGC	TGC	GGC	TGC	TGT	AAG	GTC	TGC	GCC	AAG	CAG	CTC	AAC	GAG	GAC	TGC	192
Gly	Cys	Gly	Cys	Cys	Lys	Val	Cys	Ala	Lys	Gln	Leu	Asn	Glu	Asp	Cys	
	50					55				60						
AGC	AAA	ACG	CAG	CCC	TGC	GAC	CAC	ACC	AAG	GGG	CTG	GAA	TGC	AAC	TTC	240
Ser	Lys	Thr	Gln	Pro	Cys	Asp	His	Thr	Lys	Gly	Leu	Glu	Cys	Asn	Phe	
	65				70					75					80	
GGC	GCC	AGC	TCC	ACC	GCT	CTG	AAG	GGG	ATC	TGC	AGA	GCT	CAG	TCA	GAG	288
Gly	Ala	Ser	Ser	Thr	Ala	Leu	Lys	Gly	Ile	Cys	Arg	Ala	Gln	Ser	Glu	
				85					90					95		
GGC	AGA	CCC	TGT	GAA	TAT	AAC	TCC	AGA	ATC	TAC	CAA	AAC	GGG	GAA	AGT	336
Gly	Arg	Pro	Cys	Glu	Tyr	Asn	Ser	Arg	Ile	Tyr	Gln	Asn	Gly	Glu	Ser	
			100					105					110			
TTC	CAG	CCC	AAC	TGT	AAA	CAT	CAG	TGC	ACA	TGT	ATT	GAT	GGC	GCC	GTG	384
Phe	Gln	Pro	Asn	Cys	Lys	His	Gln	Cys	Thr	Cys	Ile	Asp	Gly	Ala	Val	
		115					120					125				
GGC	TGC	ATT	CCT	CTG	TGT	CCC	CAA	GAA	CTA	TCT	CTC	CCC	AAC	TTG	GGC	432
Gly	Cys	Ile	Pro	Leu	Cys	Pro	Gln	Glu	Leu	Ser	Leu	Pro	Asn	Leu	Gly	
	130					135					140					
TGT	CCC	AAC	CCT	CGG	CTG	GTC	AAA	GTT	ACC	GGG	CAG	TGC	TGC	GAG	GAG	480
Cys	Pro	Asn	Pro	Arg	Leu	Val	Lys	Val	Thr	Gly	Gln	Cys	Cys	Glu	Glu	
	145				150					155					160	
TGG	GTC	TGT	GAC	GAG	GAT	AGT	ATC	AAG	GAC	CCC	ATG	GAG	GAC	CAG	GAC	528
Trp	Val	Cys	Asp	Glu	Asp	Ser	Ile	Lys	Asp	Pro	Met	Glu	Asp	Gln	Asp	
				165					170					175		
GGC	CTC	CTT	GGC	AAG	GAG	CTG	GGA	TTC	GAT	GCC	TCC	GAG	GTG	GAG	TTG	576
Gly	Leu	Leu	Gly	Lys	Glu	Leu	Gly	Phe	Asp	Ala	Ser	Glu	Val	Glu	Leu	
			180					185					190			
ACG	AGA	AAC	AAT	GAA	TTG	ATT	GCA	GTT	GGA	AAA	GGC	AGC	TCA	CTG	AAG	624
Thr	Arg	Asn	Asn	Glu	Leu	Ile	Ala	Val	Gly	Lys	Gly	Ser	Ser	Leu	Lys	
		195					200					205				
CGG	CTC	CCT	GTT	TTT	GGA	ATG	GAG	CCT	CGC	ATC	CTA	TAC	AAC	CCT	TTA	672
Arg	Leu	Pro	Val	Phe	Gly	Met	Glu	Pro	Arg	Ile	Leu	Tyr	Asn	Pro	Leu	
	210					215					220					
CAA	GGC	CAG	AAA	TGT	ATT	GTT	CAA	ACA	ACT	TCA	TGG	TCC	CAG	TGC	TCA	720
Gln	Gly	Gln	Lys	Cys	Ile	Val	Gln	Thr	Thr	Ser	Trp	Ser	Gln	Cys	Ser	
	225				230					235					240	
AAG	ACC	TGT	GGA	ACT	GGT	ATC	TCC	ACA	CGA	GTT	ACC	AAT	GAC	AAC	CCT	768
Lys	Thr	Cys	Gly	Thr	Gly	Ile	Ser	Thr	Arg	Val	Thr	Asn	Asp	Asn	Pro	
				245					250					255		

GAG TGC CGC CTT GTG AAA GAA ACC CGG ATT TGT GAG GTG CGG CCT TGT 816
 Glu Cys Arg Leu Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys
 260 265 270
 GGA CAG CCA GTG TAC AGC AGC CTG AAA AAG GGC AAG AAA TGC AGC AAG 864
 Gly Gln Pro Val Tyr Ser Ser Leu Lys Lys Gly Lys Lys Cys Ser Lys
 275 280 285
 ACC AAG AAA TCC CCC GAA CCA GTC AGG TTT ACT TAC GCT GGA TGT TTG 912
 Thr Lys Lys Ser Pro Glu Pro Val Arg Phe Thr Tyr Ala Gly Cys Leu
 290 295 300
 AGT GTG AAG AAA TAC CGG CCC AAG TAC TGC GGT TCC TGC GTG GAC GGC 960
 Ser Val Lys Lys Tyr Arg Pro Lys Tyr Cys Gly Ser Cys Val Asp Gly
 305 310 315 320
 CGA TGC TGC ACG CCC CAG CTG ACC AGG ACT GTG AAG ATG CGG TTC CGC 1008
 Arg Cys Cys Thr Pro Gln Leu Thr Arg Thr Val Lys Met Arg Phe Arg
 325 330 335
 TGC GAA GAT GGG GAG ACA TTT TCC AAG AAC GTC ATG ATG ATC CAG TCC 1056
 Cys Glu Asp Gly Glu Thr Phe Ser Lys Asn Val Met Met Ile Gln Ser
 340 345 350
 TGC AAA TGC AAC TAC AAC TGC CCG CAT GCC AAT GAA GCA GCG TTT CCC 1104
 Cys Lys Cys Asn Tyr Asn Cys Pro His Ala Asn Glu Ala Ala Phe Pro
 355 360 365
 TTC TAC AGG CTG TTC AAT GAC ATT CAC AAA TTT AGG GAC TAA 1146
 Phe Tyr Arg Leu Phe Asn Asp Ile His Lys Phe Arg Asp
 370 375 380

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 381 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Ser Arg Ile Ala Arg Ala Leu Ala Leu Val Val Thr Leu Leu
 1 5 10 15
 His Leu Thr Arg Leu Ala Leu Ser Thr Cys Pro Ala Ala Cys His Cys
 20 25 30
 Pro Leu Glu Ala Pro Lys Cys Ala Pro Gly Val Gly Leu Val Arg Asp
 35 40 45
 Gly Cys Gly Cys Cys Lys Val Cys Ala Lys Gln Leu Asn Glu Asp Cys
 50 55 60
 Ser Lys Thr Gln Pro Cys Asp His Thr Lys Gly Leu Glu Cys Asn Phe
 65 70 75 80
 Gly Ala Ser Ser Thr Ala Leu Lys Gly Ile Cys Arg Ala Gln Ser Glu
 85 90 95
 Gly Arg Pro Cys Glu Tyr Asn Ser Arg Ile Tyr Gln Asn Gly Glu Ser
 100 105 110
 Phe Gln Pro Asn Cys Lys His Gln Cys Thr Cys Ile Asp Gly Ala Val
 115 120 125

Gly Cys Ile Pro Leu Cys Pro Gln Glu Leu Ser Leu Pro Asn Leu Gly
 130 135 140
 Cys Pro Asn Pro Arg Leu Val Lys Val Thr Gly Gln Cys Cys Glu Glu
 145 150 155 160
 Trp Val Cys Asp Glu Asp Ser Ile Lys Asp Pro Met Glu Asp Gln Asp
 165 170 175
 Gly Leu Leu Gly Lys Glu Leu Gly Phe Asp Ala Ser Glu Val Glu Leu
 180 185 190
 Thr Arg Asn Asn Glu Leu Ile Ala Val Gly Lys Gly Ser Ser Leu Lys
 195 200 205
 Arg Leu Pro Val Phe Gly Met Glu Pro Arg Ile Leu Tyr Asn Pro Leu
 210 215 220
 Gln Gly Gln Lys Cys Ile Val Gln Thr Thr Ser Trp Ser Gln Cys Ser
 225 230 235 240
 Lys Thr Cys Gly Thr Gly Ile Ser Thr Arg Val Thr Asn Asp Asn Pro
 245 250 255
 Glu Cys Arg Leu Val Lys Glu Thr Arg Ile Cys Glu Val Arg Pro Cys
 260 265 270
 Gly Gln Pro Val Tyr Ser Ser Leu Lys Lys Gly Lys Lys Cys Ser Lys
 275 280 285
 Thr Lys Lys Ser Pro Glu Pro Val Arg Phe Thr Tyr Ala Gly Cys Leu
 290 295 300
 Ser Val Lys Lys Tyr Arg Pro Lys Tyr Cys Gly Ser Cys Val Asp Gly
 305 310 315 320
 Arg Cys Cys Thr Pro Gln Leu Thr Arg Thr Val Lys Met Arg Phe Arg
 325 330 335
 Cys Glu Asp Gly Glu Thr Phe Ser Lys Asn Val Met Met Ile Gln Ser
 340 345 350
 Cys Lys Cys Asn Tyr Asn Cys Pro His Ala Asn Glu Ala Ala Phe Pro
 355 360 365
 Phe Tyr Arg Leu Phe Asn Asp Ile His Lys Phe Arg Asp
 370 375 380

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCGGGATCC TGC GCGACAC AATGAGCT

28

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGCGGGTACC AGGTAGCATT TAGTCCCTAA

30

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAAGGATCCA CAATGAGCTC CCGAATC

27

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 58 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGCTCTAGAT TAAGCGTAGT CTGGGACGTC GTATGGGTAT TGGAACAGCC TGTAGAAG

58